De novo design of symmetric ferredoxins that shuttle electrons in vivo

Andrew C. Mutter[a], Alexei M. Tyryshkin[a], Ian J. Campbell[b], Saroj Poudel[a,c], George N. Bennett[d,e], Jonathan J. Silberg[d,e], Vikas Nanda[f,1], and Paul G. Falkowski[a,g,1]

[a]Environmental Biophysics and Molecular Ecology Program, Department of Marine and Coastal Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901; [b]Department of BioSciences, Rice University, Houston, TX; [c]Institute of Earth, Ocean, and Atmospheric Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901; [d]Department of Bioengineering, Rice University, Houston, TX; [e]Department of Chemical & Biomolecular Engineering, Rice University, Houston, TX; [f]Center for Advanced Biotechnology and Medicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; and [g]Department of Earth and Planetary Sciences, Rutgers, The State University of New Jersey, Piscataway, NJ 08854

A symmetric origin for bacterial ferredoxins was first proposed over 50 y ago, yet, to date, no functional symmetric molecule has been constructed. It is hypothesized that extant proteins have drifted from their symmetric roots via gene duplication followed by mutations. Phylogenetic analyses of extant ferredoxins support the independent evolution of N- and C-terminal sequences, thereby allowing consensus-based design of symmetric 4Fe-4S molecules. All designs bind two [4Fe-4S] clusters and exhibit strongly reducing midpoint potentials ranging from −405 to −515 mV. One of these constructs efficiently shuttles electrons through a designed metabolic pathway in *Escherichia coli*. These finding establish that ferredoxins consisting of a symmetric core can be used as a platform to design novel electron transfer carriers for in vivo applications. Outer-shell asymmetry increases sequence space without compromising electron transfer functionality.

consensus design | [4Fe-4S] clusters | bacterial ferredoxin | protein evolution | electron transfer

Biological electron transfer reactions are the source of energy for life. By definition these reactions are catalyzed by a set of oxidoreductases (EC1) which are biochemically classified by function rather than evolutionary heritage. How these enzymes originated and evolved remains one of the most enigmatic questions in science. Understanding their heritage potentially brings us closer to understanding the origins of life.

Geochemical evidence strongly suggests that for the first half (ca. 2.5 billion y) of Earth's history, oxygen concentrations were less than 0.0001% of the atmospheric volume (1). In that state, soluble, ferrous iron and sulfides would have been abundant in the oceans and would have provided an ideal environment for the biogenesis of iron-sulfur clusters in simple biomolecules (2). We chose to focus on simple ancient, ubiquitous iron-sulfur (FeS) binding proteins: the ferredoxins. We hypothesize that by studying the evolution and biochemical properties of ferredoxins, we can discover fundamental design principles for these molecules and "reverse engineer" the design of simple synthetic oxidoreductases that can support ancient metabolic processes in cells.

Ferredoxins contain FeS clusters of various stoichiometries: [2Fe-2S], [3Fe-4S], and [4Fe-4S], coordinated predominantly by cysteine and histidine first-shell ligands (3, 4). The bacterial ferredoxin fold has a $(\beta\text{-}\alpha\text{-}\beta)_2$ topology coordinating two [4Fe-4S] clusters (Fig. 1A), with pseudo-C2 symmetry, intimating a symmetric ancestor arose from gene duplication of a single $\beta\text{-}\alpha\text{-}\beta$ domain. Over time, bacterial ferredoxins have diverged from sequence symmetry (5–7).

The asymmetric sequences either arose from functional selection for a cellular role (e.g., coevolution with docking sites to electron donors or acceptors), or through independent, neutral evolution of the N- and C-terminal halves. Phylogenetic studies of extant ferredoxins show a trend toward more N- and C-terminal sequence identity with modern variants having 27% while more ancient ferredoxins having up to 59% (8). In addition to serving as monomeric soluble electron carriers, the ferredoxin

fold has been incorporated into larger electron transport chains, and used as the starting scaffold for complex metal centers (3, 4). This diversity of functional niches makes ferredoxins an attractive platform for de novo design.

Consensus-based sequence design is a powerful tool for probing the evolution of large proteins built up through the repetition of smaller precursors (9). Repeat proteins are often resilient to reshuffling or chimeric combinations of symmetric elements, allowing experimental insight into evolution and protein design strategies (9, 10). Utilization of symmetrical elements and consensus-based protein design successfully has created stable designs that probe evolutionary characteristics of sequences and structure (11, 12). Symmetry has also been a strong design element for reducing complexity in model metal binding peptides (13, 14). We consider ferredoxins as containing repeat elements and explore the extent to which symmetry dictates cluster incorporation and redox function.

Rational, consensus, and de novo design strategies have all been used to build synthetic FeS proteins. Attempts to simplify ferredoxins into halves (15) and further into small consensus sequence-derived peptides produced designs with FeS binding capacity (16). De novo designs incorporated [4Fe-4S] clusters into loops between helical elements (17, 18) or in the interior of helical bundles (19, 20). One of the smallest redox-stable designs

Significance

Early life is thought to have evolved from simple building blocks that were propagated through gene duplication events. A classic example is the small soluble iron-sulfur containing protein, bacterial ferredoxin, which is an asymmetric dimer, an essential component of many extant electron transfer chains and has ancient origins. To probe the theoretical gene duplication origins of bacterial ferredoxins, we designed a series of synthetic symmetric constructs. All designs bound two iron-sulfur clusters and were able to support electron transfer between a pair of oxidoreductases in vivo in *Escherichia coli*. Our results strongly suggest that simple, symmetric ancestral proteins probably evolved early in Earth's history and can be engineered to facilitate functional electron transfer in synthetic metabolic pathways.

BIOCHEMISTRY

Fig. 1. Structural symmetry and consensus designs. (*A*) Color blocked diagram of bacterial ferredoxin's asymmetric sequence accompanied by a cartoon representation of PDB ID 1FDN with axis of symmetry highlighted by a dot, N- and C-terminal colored magenta and cyan, respectively. (*B*) Structural alignment of N and C structures showing structural symmetry of the parent protein. (*C*) Unrooted maximum-likelihood phylogenetic tree of N-terminal motif (red) and C-terminal motif (blue) of ferredoxin. Bootstrap value of each node is >54. (*D*) Color block diagrams of asymmetric vs. symmetric designs with consensus sequence aligned above each diagram with negative and positive residues colored red and blue, respectively, along with highlighting for symmetric core residues (gray) and variable outer-shell residues (orange), * and $ indicate cluster binding.

to date was constructed from both l- and d-amino acids to produce a minimal 12-residue peptide (21).

Here, we characterized multiple consensus sequence designs with a bacterial ferredoxin fold. All symmetric and asymmetric designs bound two [4Fe-4S] clusters and were capable of reversible electron transfer, experimentally validating that a putative symmetric ferredoxin ancestor could have functioned as a mobile electron carrier. The high success rate of our designs shows that ferredoxins are a highly designable scaffold for creating tunable electron transfer proteins, where limited constraints on sequence symmetry enlarge the functional sequence footprint of redox activity. Within these constraints our constructs allow for in vivo electron cycling in designed *Escherichia coli* cellular pathways.

Results

For a consensus design of symmetric ferredoxins to successfully fold and bind two [4Fe-4S] clusters, we posited that the N- and C-terminal halves would have evolved independently. If true, symmetric constructs containing duplications of either motif would likely be functional. We used phylogenetic analysis to observe how evolution constrained symmetry between N- and C-terminal halves and evolutionary coupling analysis to identify the extent of coevolution between these two domains.

Coevolution between N- and C-Terminal Motifs. We divided a set of natural ferredoxins into N- and C-terminal fragments and evaluated similarity across N vs. N, C vs. C, and N vs. C alignments. In the case of independent evolution we expected the N and C motifs to cluster on distinct branches of a phylogenetic tree. Maximum-likelihood phylogenetic tree analysis (Fig. 1*C*) showed distinct clades for both N- (highlighted red in Fig. 1*C*) and C-terminal (highlighted blue in Fig. 1*C*) which was consistent with independent evolution of the two motifs. Each ferredoxin species is represented twice on the tree, once each for its N- and C-terminal halves. This supports largely independent evolution of variable positions in the ferredoxin fold after an ancestral duplication event.

An independent approach for identifying coevolving intra- and interresidue contacts from sequence information is evolutionary coupling (EC) analysis which compares coupled residue mutations across protein evolutionary families. The strongest couplings generally arise from residues that are in physical contact in the protein structure. In instances where this is not the case, the coupling may instead be due either to allosteric motions or functional dependence. Using the EVcouplings program (22) 12

pairs of residues were identified with EC scores >0.2. Of these, only two pairs, positions 24 and 36 as well as positions 20 and 51, showed coupling between domains (*SI Appendix*, Fig. S2). Both pairs were not in physical contact (*SI Appendix*, Fig. S2, *Inset*), suggesting functional rather than structural constraints on coevolution. A significant fraction of residues participating in contacts between the two domains were highly conserved across the multiple sequence alignment, and therefore lacked sufficient mutual information to specify coevolutionary constraints.

Consensus Designs. Designed ferredoxins were derived from two datasets resulting in two families with three members. An asymmetric design was constructed using the consensus sequence directly (*SI Appendix*, Fig. S1), along with two symmetric designs based on duplications of the N- or C-terminal consensus sequence halves (Fig. 1*D*). The first family of sequences (**ANC**, **ANN**, and **ACC**) were derived from 150 ferredoxin homologs from the Uniref-90 database (23). This included embedded ferredoxin domains from large multicluster-containing proteins, which led us to anticipate these designs may suffer from limited solubility due to consensus sequence preferences driven by interior hydrophobic positions. To address this concern, a second set of designs (**SNC**, **SNN**, and **SCC**) was generated based on consensus sequences derived from a dataset consisting of 17 soluble ferredoxin sequences (*SI Appendix*, Table S2). **ANC** and **SNC** shared only 58.2% sequence identity, motivating characterization of both sets of designs.

Within **ANC** and **SNC**, there was 50% identity between the N- and C-terminal motifs. Residues whose position and identity matched symmetry of the overall fold specified an inner shell that coordinated metal and formed a hydrophobic core within and between the two domains (*SI Appendix*, Fig. S3). The remaining positions formed a less-conserved outer shell that accounted for most of the sequence variability across all designs. The outer shell of the **SNC** and derived designs had a high frequency of acidic residues (11 aspartate and glutamate residues versus one lysine), resulting in isoelectric points (pIs) <4. In contrast, **ANC** and associated designs had pIs from 6.9 to 8.2 stemming from the more balanced 7 aspartate and glutamate residues versus 10 lysine and arginine residues. This reflects differences in the makeup of soluble vs. embedded ferredoxins, where the net-negative charge of soluble ferredoxins may facilitate solubility at physiological pH and/or promote interactions with partners in an electron transfer pathway (24).

Iron Incorporation. Proteins were aerobically expressed as described in *Materials and Methods*. All synthetic ferredoxins expressed well (20 mg/L from SDS/PAGE) using autoinduction and produced a brown cell pellet when harvested. Stable in vivo cluster incorporation was inferred from persistence of brown color during initial purification steps under anaerobic conditions, versus loss of color in the presence of oxygen. Initially, purified proteins did not fully bind FeS clusters as inferred from size exclusion elution profiles. Apo ferredoxins eluted as a higher molecular weight broad peak that sharpened if preincubated with 5 mM dithiothreitol (DTT) (*SI Appendix*, Fig. S4A). To improve the yield and consistency of FeS incorporation, we reconstituted the purified proteins with iron and sulfur in vitro.

In vitro FeS incorporation from an inorganic mixture of precursors was performed following established protocols (19). Purified reconstituted proteins show absorbance peaking at 415 nm typical of a [4Fe-4S] cluster (25) Addition of dithionite eliminated this absorbance as expected (Fig. 2). Cluster incorporation reaction was confirmed by size exclusion chromatography with reconstituted proteins eluting as one peak slightly smaller than apo protein when eluted with 1 mM DTT in all buffers (*SI Appendix*, Fig. S4A). All designs showed successful incorporation of [4Fe-4S] clusters with a stoichiometry of two clusters per protein, as determined by comparing cluster yield electron paramagnetic resonance (EPR) spin counting to peptide concentration by UV-visible absorbance.

EPR Characterization of Consensus Designs. Relevant spectroscopic evidence of iron sulfur cluster formation can be discerned with EPR (25). Each type of iron sulfur cluster has an unique g-factor signature, allowing screening of in vitro FeS reconstitution products, and probing of the local cluster environments (Fig. 3). For example, reduced [4Fe-4S]$^{1+}$ and oxidized [3Fe-4S]$^{1+}$ clusters are expected to have an unpaired spin S = 1/2, thus being detectable by EPR. The oxidized samples for all our designs showed no EPR signal (Fig. 3), thus ruling out a [3Fe-4S]$^{1+}$ cluster formation. After reduction of [4Fe-4S]$^{2+}$ to [4Fe-4S]$^{1+}$ with addition of an excess of dithionite (20 mM), the samples showed complex spectra with several prominent features such as distinctive shoulders (Fig. 3 A and B), indicative of successful incorporation of two [4Fe-4S] clusters per design. The two reduced [4Fe-4S]$^{1+}$ strongly interact with each other through spin exchange and dipolar interactions producing the spectra as observed (26–28). All of the designs showed similar EPR spectra except **ACC** (Fig. 3). The spectra for **ACC** revealed clear differences in the

shape of the center signal and the relative intensity of the side shoulders (Fig. 3 A and B). These differences between **ACC** and others are discussed in the next section.

Incorporation of two [4Fe-4S] clusters per design, one in each of the two sites was confirmed by identifying dipolar (D) and exchange (J) interactions in the EPR spectra. The spectra as observed, with a narrow strong central feature and broad weak shoulders, can only arise when spin–spin interaction between two [4Fe-4S] clusters is much stronger than the Zeeman detuning of their spins: $|J \pm D| \gg \Delta g \beta B_0$, where Δg is the orientation-dependent difference in g-factors of two clusters, β is Bohr magneton, B_0 is the applied magnetic field (29, 30). In this case, the width of the center feature should be approximately equal to $|3/2 \cdot D|$, and the splitting between the side shoulders should be equal to $|2 \cdot J + D/2|$. The dipolar coupling D = 200 ± 50 MHz and the exchange coupling $|J|$ = 800 ± 200 MHz can be estimated from the spectra in Fig. 3. These estimates are valid for all six designs, including **SCC** and **ACC**. The estimated D = 200 MHz corresponds to a distance of 8 ± 1 Å between the nearest Fe ions of the coupled [4Fe-4S]$^{1+}$ clusters. This correlates with detailed spectral simulations on *Clostridium pasteurianum* (29), and matches the 8.7 Å separation observed in the high-resolution soluble ferredoxin structure from *Clostridium acidurici* (2FDN) (31).

Redox Function. Redox titrations of the designed proteins were performed using the optically transparent thin layer electrochemical (OTTLE) cell technique (32) using a commercially available honeycomb gold electrode. All designs showed fully reversible oxidation and reduction as shown by absorption change at 430 nm (Fig. 4A). The **SNC** family of proteins exhibited midpoint potentials between −498 and −512 mV, while those of the **ANC** constructs were between −403 and −486 mV, representing a span of 2.5 kcal in energy. Midpoint potentials correlate with the pIs of the designs, with more acidic proteins acting as stronger reducers. The variation in pI is largely due to titratable outer-shell residues on the protein surface, so this phenomenon may be largely understood as through-space electrostatic stabilization of the [4Fe-4S]$^{2+}$ state by the net-negative charge on **SNC**, **SNN**, and **SCC** at pH 8.0. This observation is consistent with similar observations in other designed redox active proteins (33). A larger variation in midpoint potential is observed for the **ANC** series with **ACC** and **ANN** potentials bracketing that of the asymmetric design. For these designs, the calculated pI is close to the experimental pH 8, where significant variation in net charge across designs may be expected.

Titration of a Core Histidine. While most of the asymmetric positions in the outer shell are solvent exposed, one exception is a histidine at position 30 (H30) in **ANC**. This buried histidine packs directly against one of the two [4Fe-4S] clusters in structural modeling (Fig. 4). This cluster-proximal histidine is commonly found in NiFe hydrogenases (34) and thermostable ferredoxins (15, 35). Titration of this histidine results in a significant shift in midpoint potential for these naturally occurring ferredoxins (36).

To evaluate the proximal effect from H30, we measured the pH dependence of redox potentials in the **ANC** family (Fig. 4B). Raising the pH from 8.0 to 10.0 likely deprotonates this residue, making **ANC** more reducing by −12 mV. In the symmetric **ACC**, a second histidine is introduced at position 1, resulting in nearly double the shift in midpoint potential of 33 mV, corresponding to ~0.75 kcal/mol for moving one electron. **ANN** (Fig. 4B), which has tyrosines as well as **SCC** (*SI Appendix*, Fig. S6) at these positions, shows a minimal positive shift in midpoint potential upon increasing pH.

The unusual EPR spectrum of **ACC** may be induced by distortions in inner-shell interactions by H30 and H2's proximity to the [4Fe-4S] clusters. Despite differences in EPR spectra, a similar spatial arrangement of the [4Fe-4S]$^{1+}$ clusters in **ACC** compared with the other designs is likely (Fig. 3). We speculate that differences arise from different angular orientations of the cluster

Fig. 2. UV-Vis spectra of six consensus-designed proteins after in vitro [4Fe-4S] cluster incorporation in oxidized form. *Inset* shows the spectra for oxidized (solid) and dithionite-reduced (dashed) **SCC**.
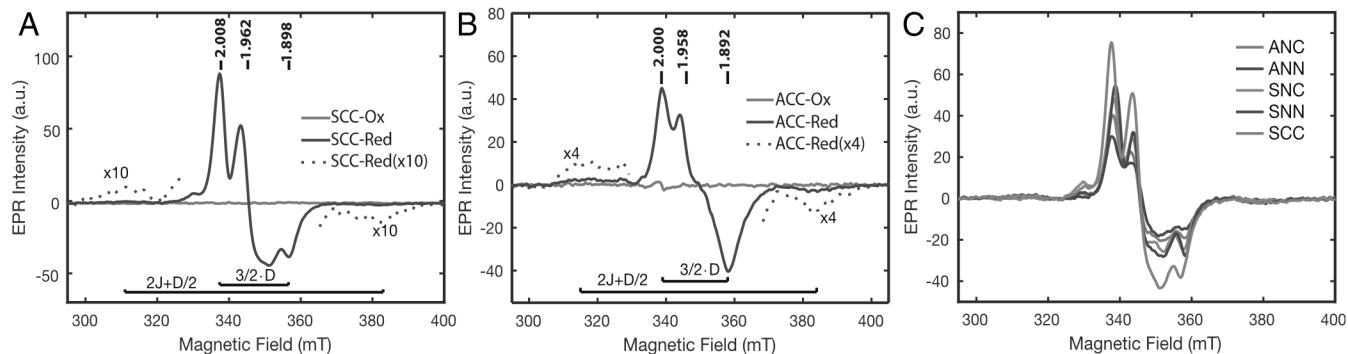
Fig. 3. X-band EPR spectra of six designed [4Fe-4S] proteins at 10K: (*A*) **SCC** and (*B*) **ACC** in their oxidized (blue) and fully reduced (red) forms. (*C*) Remaining proteins (**ANC**, **ANN**, **SNC**, and **SNN**) in their fully reduced forms shown in comparison with **SCC**. In *A* and *B*, the wings of the **ACC** and **SCC** spectra are magnified by 4× and 10×, respectively, to accentuate broad symmetric shoulders on low- and high-field sides of the spectra (dashed traces). The characteristic g-factor values are marked with vertical lines atop the **ACC** and **SCC** spectra. The horizontal bars shown at the bottom in *A* and *B* identify important line splittings in the spectra.

g-factors. For a mixed-valence [4Fe-4S]$^{1+}$ cluster, orientation of its g-factor tensor depends on valence-charge distribution among four Fe ions constituting the cluster (37). Two Fe ions in the cluster maintain a higher charge, forming a mixed-valence pair (Fe$^{2.5+}$-Fe$^{2.5+}$) and the two other Fe ions maintain a lower charge, forming a ferrous pair (Fe^{2+}-Fe^{2+}). The principal g$_1$ axis of g-factor tensors is known to be oriented perpendicular to both vectors defining the mixed-valence and ferrous pairs (37). The proximity of H30 histidine to one of the [4Fe-4S]$^{1+}$ clusters in **ACC** may alter its valence-charge distribution compared with the remaining designs. Future studies will delineate the full g-tensor approximations, allowing detailed spectral simulations of this design variant.

Cellular Electron Transfer. To investigate whether the **ANC** and **SNC** family constructs support electron transfer in cells, we evaluated their ability to transfer electrons from *Zea mays* Ferredoxin NADPH Reductase (FNR) to Sulfite I Reductase (SIR) in *E. coli* EW11 (38). This strain presents growth defects on minimal medium containing sulfate as a sulfur source unless it expresses a ferredoxin that can efficiently transfer electrons from FNR to SIR. For these measurements, we evaluated the effects of each synthetic ferredoxin on cell growth under microaerophilic (0.2% O$_2$) conditions. These conditions were chosen to minimize oxidative damage to the [4Fe-4S] clusters, while enabling low levels of respiration in *E. coli*.

Because the cellular burden of each synthetic ferredoxin is unknown, we evaluated their growth complementation over a range of expression levels using an anhydrotetracycline (aTc)-

inducible promoter. FNR and SIR were constitutively expressed at levels that enable efficient coupling with plant-type [2Fe-2S] ferredoxins (39). With synthetic ferredoxins from both families, growth was enhanced by aTc (Fig. 5), albeit to differing extents. Cells expressing **ANN** presented significant increases in exponential growth rate upon aTc exposure at the highest level tested (200 ng/mL), but this gain was smaller than that observed with the **SNC** and **SCC** ferredoxins, both of which showed significant growth differences at lower aTc concentrations. The low level of growth in the absence of aTc is thought to have arisen because the cells used to initiate our growth assay contained a small amount of residual sulfide. To remove this sulfur, we washed the cells from our starter cultures before inoculation and initiation of the growth assay. With more rigorous washing, only **SCC** showed any growth upon induction (*SI Appendix*, Fig. S8). This approach provides a tool for modulating the stringency of selection to identify designs with a broad range of electron shuttling activities.

In all cases, the growth enhancement was smaller than that observed with a native [2Fe-2S] ferredoxin having a more positive midpoint potential (−336 mV) that more closely matches the midpoint potential of the FNR electron donor (−337 mV) used in the cellular assay (39). The growth enhancements of each synthetic ferredoxin did not correlate with their midpoint potential offset from FNR. Two of the **SNC** family ferredoxins, **SNC** and **SCC**, which have the lowest potentials, supported growth enhancement to the greatest extent and displayed the greatest differences in lag time between uninduced and induced strains (*SI Appendix*, Table S1). The 160- to 170-mV uphill barrier between FNR and designed
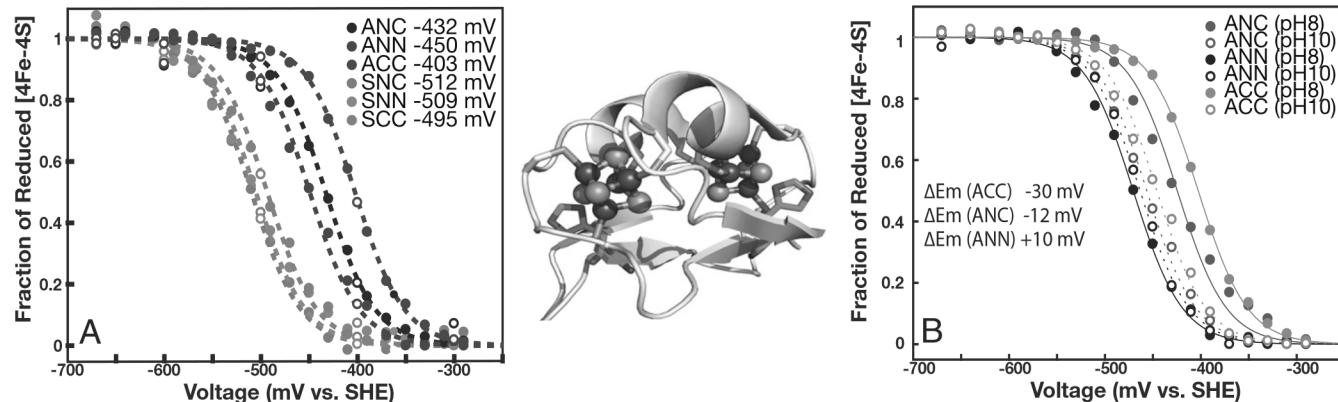
Fig. 4. Redox titrations of consensus design proteins fit to Nernst equation with *n* = 1. Closed circles are reduction points, open circles oxidation points. (*A*) All constructs titrations performed at ph 8. (*B*) pH-dependent shifts in midpoint potential of **ANC** and **ACC**. (*Inset*) Model of PDB ID 1FDN showing positions of H1 and H30 proximal to the [4Fe-4S] clusters in **ACC**. All redox titrations were performed on concentrated samples of ~2 mg/mL protein.
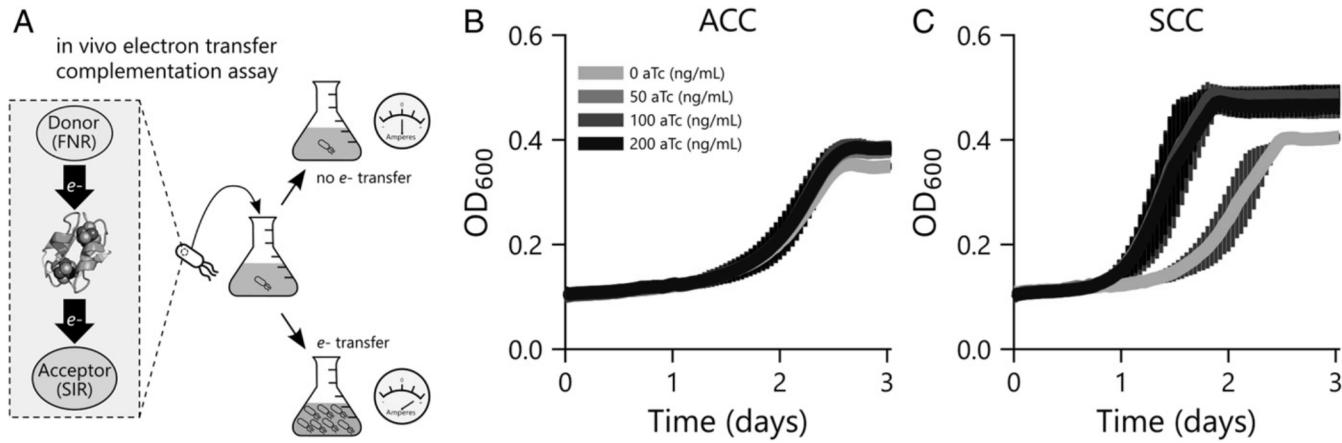
Fig. 5. Synthetic ferredoxin-enhanced growth in an *E. coli* auxotroph. (*A*) Schematic of ferredoxin complementation in auxotrophic *E. coli* EW11 grown on minimal media containing sulfate as sulfur source by enabling electron transfer from ferredoxin NADPH reductase (FNR) to sulfite reductase (SIR). The production of sulfide enables cell growth. (*B* and *C*) Growth of sulfide auxotroph with ferredoxin constructs from the ANC or SNC families (*n* = 3). Ferredoxin expression was controlled by an aTc-inducible promoter. Gray error bars depict SD.

ferredoxins would likely result in minimal electron transfer, indicating that binding may shift either or both donor and acceptor midpoint potentials to the extent that transfer is energetically feasible. Acidic outer-shell designs **SNC** and **SCC** may interact tightly with the highly basic docking sites of the two *Z. mays* partner proteins (*SI Appendix*, Fig. S7). Intermolecular charge pair interactions at the FNR–ferredoxin and SIR–ferredoxin interface may perturb local electrostatics to a sufficient extent to drive midpoint potential shifts.

Discussion

Phylogenetic and evolutionary coupling analysis indicates that the two halves of bacterial ferredoxins have a conserved inner shell that is largely symmetric in sequence and an outer shell where the N- and C-terminal regions can evolve independently. This finding is consistent with the ubiquity of the ferredoxin fold in nature (3, 4, 7, 40), where the high evolvability provides a large functional footprint in sequence space. This large sequence space is also of great benefit in the design of synthetic electron shuttles with target properties (39, 41). The N- and C-terminal outer-shell residues can be independently sampled to tune midpoint potential or develop binding sites for donors and acceptors in an electron transfer circuit. The bacterial ferredoxin fold is likely amenable to dynamic library-based design approaches (42, 43), where fragments sharing the inner-shell residues but with varied outer shells can be recombined to create libraries of assemblies and screened.

Sampling of large libraries and directed-evolution methods are greatly simplified by the expression of these constructs in vivo. The stringency of selection may be tunable by varying the amount of reduced sulfur in the growth medium, allowing optimization of designs with minimal initial activity. Roy and colleagues have shown in vitro electron transfer from a de novo designed ferredoxin to a natural cytochrome *c* acceptor is possible (44). In vivo realization of an electron transfer pathway allows for concurrent optimization of both redox function and binding partner affinity and specificity. Using designable folds such as that of bacterial ferredoxin enhance the likelihood that high-throughput selections imposed on naïve proteome libraries (45) will lead to successful hits.

Tuning of midpoint potential seems to obey simple rules. If the inner-shell chassis of ferredoxin is maintained, outer-shell residues can be rationally designed to reflect a spectrum of charge states, with more anionic designs favoring a strongly reducing ferredoxin. More dramatic changes and stimulus-responsive redox proteins may be introduced by modulating inner-shell residues such as the proximal histidine studied here. Synthetic ferredoxins can be used to study the functional role of natural inner-shell variants in stability, redox, and catalysis.

The same features that have promoted ferredoxin evolution make it an attractive platform for design and metabolic engineering. Ferredoxin's ancient evolutionary past is linked to the success of a short robust sequence that potentially permits modular construction of more complex oxidoreductases. Currently, we are limited in our ability to construct energy-conserving metabolic pathways that avoid global redox imbalances which can limit chemical yields in metabolic engineering. Our artificial ferredoxins clearly show that improved design rules for synthetic ferredoxins can enable the creation of carriers that control and optimize electron flow between diverse oxidoreductases. Such ferredoxins will be useful in synthetic biology for creating energy-conserving pathways and living electronic sensors.

Materials and Methods

Consensus Design. *C. acidurici* ferredoxin (PDB ID: 2FDN) (31) was used as a bait sequence in the ConSurf server (46) to identify homologous ferredoxin sequences using the default settings. Two families of consensus sequences were generated: one directly from the ConSurf multiple sequence alignment involving 150 structures and from a second nonredundant set of 17 structurally determined soluble ferredoxins (*SI Appendix*, Table S2). These parent consensus sequences were sequence asymmetric and were then used to generate symmetric child sequences, either by repeating the first (N terminal) or last (C terminal) 26 amino acids of the parent sequence, separated by the three amino acid linker of the parent sequence (Fig. 1*D*).

Phylogenetic Analysis. The amino acid sequences of a nonredundant set of structurally determined soluble bacterial ferredoxins (*SI Appendix*, Table S2) were used as bait sequences to BLAST against the NCBI complete genome database. The top 10 homologs of each bait sequence were extracted and identical sequences were removed for downstream analysis. The resulting 347 sequences were aligned using the MAFFT server (47) with default settings. Positions with >10% gaps in the alignment were eliminated. This removed large insertions, resulting in only two contiguous stretches of sequence common to the entire alignment which were defined as the N- and C-terminal motifs. These sequences were subjected to maximum-likelihood phylogenetic reconstruction with RAxML (version 7.3) (48), specifying the LG substitution matrix (49), the PROTGAMMA option, and 1,000 bootstrap iterations. The reconstructed tree was visualized with iTOL (50).

Gene Construction and Protein Expression. Synthetic genes codon optimized for *E. coli*, were prepared by Genewiz, Inc. and cloned into pUC57. Gibson Assembly kit from NEB was used for subcloning synthetic genes behind an octa-histidine-tagged maltose-binding protein separated by a TEV protease cut site in custom expression plasmids (gift from John D. Kim, Rutgers, The State University of New Jersey, New Brunswick, NJ) following manufacturer's instructions. Synthetic ferredoxins were coexpressed with the inducible *isc-SUA-hscBA-fdx* operon using plasmid pDB1282 (gift from Dennis Dean, Virginia Tech, Blacksburg, VA) following autoinduction protocol (51). Briefly, 3 mL LB overnight cultures were used to inoculate 200 mL LB-based autoinducing

media, including both 0.2% arabinose and 0.2% lactose. Resulting recombinant proteins were purified by standard methods (*SI Appendix*).

Protein Potentiometry. UV-Vis redox titrations were performed using a Pine Research gold honeycomb electrode and AgCl reference electrode controlled via a BioLogic SP-50 potentiostat under flow of oxygen-free nitrogen in a glovebag. Full spectra were recorded on a Cary 60 spectrophotometer with 430 nm used to monitor reduction.

EPR Spectroscopy. EPR samples were prepared anaerobically at protein concentrations of 50 to 100 μM in a 100-mM Tris or 100 mM phosphate buffer with 100 mM NaCl (pH 8.0) supplemented with 15% glycerol as a cryoprotectant. Excess sodium dithionite (20 mM) was added to fully reduce all iron-sulfur clusters to $[4Fe\text{-}4S]^{1+}$ states. EPR spectra were recorded on a Bruker EPR spectrometer (E580e) operating at X-band microwave frequency. A helium-flow cryostat (Oxford ESR900) equipped with an Oxford temperature controller (ITC503) was used for cryogenic temperatures. Typical experimental conditions were: microwave frequency, 9.49 GHz; microwave power, 0.2 mW; modulation amplitude, 1 mT; and temperature, 10 K. Concentrations of reduced $[4Fe\text{-}4S]^{1+}$ clusters were determined using EPR spin counting, by comparing the measured signal intensities with an EPR standard of known weight (a $CuSO_4\cdot5H_2O$ crystal in a mineral oil).

Growth Assay. A pair of previously described plasmids was used to evaluate ferredoxin electron transfer in cells (39). *E. coli* (EW11) was cotransformed with pSAC01 and one of the pFd vectors. Cells were selected on LB-agar plates and transferred to minimal M9 complete medium (M9c) containing 34 μg/mL chloramphenicol and 100 μg/mL streptomycin (39). Cells were grown in the presence of the indicated aTc concentrations with terminal electron acceptor (6 g/L trimethylamine *N*-oxide) in a Tecan Spark plate reader at 37 °C under a 0.2% O_2 atmosphere with shaking at 300 rpm in double orbital mode. For stringent growth assays, cells were washed and resuspended once in 1 mL of selective media before growth in low-oxygen conditions. Optical density (OD) measurements were acquired every 5 min. Exponential growth rates were calculated using PMAnalyzer (52).

1. T. W. Lyons, C. T. Reinhard, N. J. Planavsky, The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307–315 (2014).
2. R. J. P. Williams, The Bakerian lecture, 1981: Natural selection of the chemical elements. *Proc. R. Soc. Lond. B Biol. Sci.* **213**, 361–397 (1981).
3. H. Beinert, Iron-sulfur proteins: Ancient structures, still full of surprises. *J. Biol. Inorg. Chem.* **5**, 2–15 (2000).
4. R. Cammack, "Iron and sulfur in the origin and evolution of biological energy conversion systems" in *Origin and Evolution of Biological Energy Conversion*, Ed H. Baltscheffsky (Wiley-VCH, 1996), pp. 43–69.
5. R. V. Eck, M. O. Dayhoff, Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* **152**, 363–366 (1966).
6. B. K. Davis, Molecular evolution before the origin of species. *Prog. Biophys. Mol. Biol.* **79**, 77–133 (2002).
7. J. Meyer, Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. *J. Biol. Inorg. Chem.* **13**, 157–170 (2008).
8. M. L. Romero Romero, A. Rabin, D. S. Tawfik, Functional proteins from short peptides: Dayhoff's hypothesis turns 50. *Angew. Chem. Int. Ed. Engl.* **55**, 15966–15971 (2016).
9. B. Höcker, Design of proteins from smaller fragments-learning from evolution. *Curr. Opin. Struct. Biol.* **27**, 56–62 (2014).
10. S. Eisenbeis, B. Höcker, Evolutionary mechanism as a template for protein engineering. *J. Pept. Sci.* **16**, 538–544 (2010).
11. A. Broom *et al.*, Modular evolution and the origins of symmetry: Reconstruction of a three-fold symmetric globular protein. *Structure* **20**, 161–171 (2012).
12. A. R. D. Voet *et al.*, Computational design of a self-assembling symmetrical β-propeller protein. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15102–15107 (2014).
13. A. Lombardi *et al.*, Miniaturized metalloproteins: Application to iron-sulfur proteins. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11922–11927 (2000).
14. A. Lombardi *et al.*, Retrostructural analysis of metalloproteins: Application to the design of a minimal model for diiron proteins. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6298–6305 (2000).
15. T.-C. Sow, M. V. Pederson, H. E. M. Christensen, B.-L. Ooi, Total synthesis of a mini-ferredoxin. *Biochem. Biophys. Res. Commun.* **223**, 360–364 (1996).
16. S. E. Mulholland, B. R. Gibney, F. Rabanal, P. L. Dutton, Determination of nonligand amino acids critical to [4Fe-4S]2+/+ assembly in ferredoxin maquettes. *Biochemistry* **38**, 10442–10448 (1999).
17. B. R. Gibney, S. E. Mulholland, F. Rabanal, P. L. Dutton, Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15041–15046 (1996).
18. C. E. Laplaza, R. H. Holm, Helix-loop-helix peptides as scaffolds for the construction of bridged metal assemblies in proteins: The spectroscopic A-cluster structure in carbon monoxide dehydrogenase. *J. Am. Chem. Soc.* **123**, 10255–10264 (2001).
19. J. Grzyb *et al.*, Empirical and computational design of iron-sulfur cluster proteins. *Biochim. Biophys. Acta* **1817**, 1256–1262 (2012).
20. A. Roy, I. Sarrou, M. D. Vaughn, A. V. Astashkin, G. Ghirlanda, De novo design of an artificial bis[4Fe-4S] binding protein. *Biochemistry* **52**, 7586–7594 (2013).
21. J. D. Kim *et al.*, Minimal heterochiral de Novo designed 4Fe-4S binding peptide capable of robust electron transfer. *J. Am. Chem. Soc.* **140**, 11210–11213 (2018).
22. D. S. Marks *et al.*, Protein 3D structure computed from evolutionary sequence variation. *PLoS One* **6**, e28766 (2011).
23. B. E. Suzek, H. Huang, P. McGarvey, R. Mazumder, C. H. Wu, UniRef: Comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288 (2007).
24. M. S. Lawrence, K. J. Phillips, D. R. Liu, Supercharging proteins can impart unusual resilience. *J. Am. Chem. Soc.* **129**, 10110–10112 (2007).
25. W. V. Sweeney, J. C. Rabinowitz, Proteins containing 4Fe-4S clusters: An overview. *Annu. Rev. Biochem.* **49**, 139–161 (1980).
26. R. Mathews, S. Charlton, R. H. Sands, G. Palmer, On the nature of the spin coupling between the iron-sulfur clusters in the eight-iron ferredoxins. *J. Biol. Chem.* **249**, 4326–4328 (1974).
27. R. C. Prince, M. W. Adams, Oxidation-reduction properties of the two Fe4S4 clusters in Clostridium pasteurianum ferredoxin. *J. Biol. Chem.* **262**, 5125–5128 (1987).

28. J. Gaillard, I. Quinkal, J. M. Moulis, Effect of replacing conserved proline residues on the EPR and NMR properties of Clostridium pasteurianum 2[4Fe-4S] ferredoxin. *Biochemistry* **32**, 9881–9887 (1993).
29. C. More *et al.*, A new approach for the structural study of metalloproteins: The quantitative analysis of intercenter magnetic interactions. *JBIC* **1**, 152–161 (1996).
30. G. Zwanenburg, P. J. Hore, EPR of spin-correlated radical pairs. Analytical treatment of selective excitation using zero-quantum coherence. *Chem. Phys. Lett.* **203**, 65–74 (1993).
31. Z. Dauter, K. S. Wilson, L. C. Sieker, J. Meyer, J.-M. Moulis, Atomic resolution (0.94 A) structure of Clostridium acidurici ferredoxin. Detailed geometry of [4Fe-4S] clusters in a protein. *Biochemistry* **36**, 16065–16073 (1997).
32. W. R. Heineman *et al.*, "Studies of biological redox systems by thin-layer electrochemical techniques" in *Electroanalytical and Spectroscopical Studies of Biological Redox Components*, K. M. Kadish, Ed. (American Chemical Society, Washington, DC, 1982), pp. 1–21.
33. D. E. Robertson *et al.*, Design and synthesis of multi-haem proteins. *Nature* **368**, 425–432 (1994).
34. L. A. Flanagan, H. S. Chidwick, J. Walton, J. W. B. Moir, A. Parkin, Conserved histidine adjacent to the proximal cluster tunes the anaerobic reductive activation of Escherichia coli membrane-bound [NiFe] hydrogenase-1. *ChemElectroChem* **5**, 855–860 (2018).
35. M. F. Perutz, H. Raidt, Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* **255**, 256–259 (1975).
36. E. T. Smith, B. A. Feinberg, Redox properties of several bacterial ferredoxins using square wave voltammetry. *J. Biol. Chem.* **265**, 14371–14376 (1990).
37. F. Moriaud, S. Gambarelli, B. Lamotte, J.-M. Mouesca, Detailed proton Q-band ENDOR study of the electron spin population distribution in the reduced [4Fe-4S] 1+ state. *J. Phys. Chem. B* **105**, 9631–9642 (2001).
38. B. Barstow *et al.*, A synthetic system links FeFe-hydrogenases to essential E. coli sulfur metabolism. *J. Biol. Eng.* **5**, 7 (2011).
39. J. T. Atkinson *et al.*, Metalloprotein switches that display chemical-dependent electron transfer in cells. *Nat. Chem. Biol.* **15**, 189–195 (2019).
40. H. Beinert, R. H. Holm, E. Münck, Iron-sulfur clusters: Nature's modular, multipurpose structures. *Science* **277**, 653–659 (1997).
41. C. M. Agapakis, P. A. Silver, Modular electron transfer circuits for synthetic biology: Insulation of an engineered biohydrogen pathway. *Bioeng. Bugs* **1**, 413–418 (2010).
42. M. A. Case, G. L. McLendon, A virtual library approach to investigate protein folding and internal packing. *J. Am. Chem. Soc.* **122**, 8089–8090 (2000).
43. S. L. Springs *et al.*, A multigeneration analysis of cytochrome b(562) redox variants: Evolutionary strategies for modulating redox potential revealed using a library approach. *Biochemistry* **41**, 4321–4328 (2002).
44. A. Roy *et al.*, A de novo designed 2[4Fe-4S] ferredoxin mimic mediates electron transfer. *J. Am. Chem. Soc.* **136**, 17343–17349 (2014).
45. A. E. Donnelly, G. S. Murphy, K. M. Digianantonio, M. H. Hecht, A de novo enzyme catalyzes a life-sustaining reaction in Escherichia coli. *Nat. Chem. Biol.* **14**, 253–255 (2018).
46. H. Ashkenazy *et al.*, ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
47. S. Kuraku, C. M. Zmasek, O. Nishimura, K. Katoh, aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **41**, W22–W28 (2013).
48. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
49. S. Q. Le, O. Gascuel, An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).
50. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
51. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
52. D. A. Cuevas, R. A. Edwards, PMAnalyzer: A new web interface for bacterial growth curve analysis. *Bioinformatics* **33**, 1905–1906 (2017).